Lecture 5 – Bivariate Random Variables

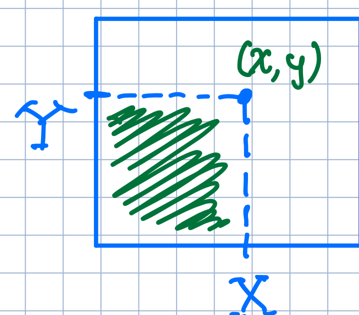
In statistical thermodynamics, we usually need to deal with functions of more than one variable. In fact, we will deal with functions of many variables. How do we deal with this in terms of random variables?

A. Joint and Marginal Distributions

To describe probabilities with more than one R.V., we need a joint distribution function. In the bivariate case the joint distribution describes how the probability depends on x and y and all interactions between x and y. In other words, we will want a distribution to describe $P(X \cap Y)$ "and" for all $X \in \mathbb{R}$ and $Y \in \mathbb{R}$. This will necessarily include all $P(X)$ and $P(Y)$.

The joint cumulative distribution function (cdf) is defined as

$$P(X \le x, Y \le y) = F_{xy}(x,y)$$



The joint distribution contains all the probability information for the "2D plane" of both x and y.

The marginal cumulative distribution function contains only the information for one of the variables. It can be obtained from the joint cdf:

$$F_x = F_{xy}(x, \infty)$$

Recall that $F(Y \le y) \to 1$ as $y \to \infty$

$$F_y = F_{xy}(\infty, y)$$

Similarly, we can define joint and marginal probability density functions

$$P((x,y)\in A) = \iint_A f_{xy}(x,y)\,dx\,dy$$ ← joint pdf

$$f_x(x) = \int_{-\infty}^{\infty} f_{xy}(x,y)\,dy$$

Recall that $\int_{-\infty}^{\infty} f_y(y)\,dy = 1$

We say we "integrate out" y.

$$f_y(y) = \int_{-\infty}^{\infty} f_{xy}(x,y)\,dx.$$

The definition of the joint cdf and pdf implies that

$$\frac{\partial^2 F_{xy}}{\partial x \partial y} = \frac{\partial^2 F_{xy}}{\partial y \partial x} = f_{xy}$$

joint cdf → joint pdf

and

$$\frac{\partial}{\partial x} F_{xy}(x,\infty) = \frac{\partial}{\partial x} F_x = f_x$$

$$\frac{\partial}{\partial y} F_{xy}(\infty,y) = \frac{\partial}{\partial y} F_y = f_y$$

marginal cdf → marginal pdf

Finally, for discrete RVs, we have joint and marginal pmfs

$$P(X=x_i, Y=y_j) = p_{xy}(x_i,y_j)$$ joint pmf

$$p_x(x_i) = \sum_{j=1}^{\infty} p_{xy}(x_i,y_j)$$ marginal pmf

sum over all y.

$$p_y(y_j) = \sum_{i=1}^{\infty} p_{xy}(x_i,y_j)$$

Example: Bivariate Gaussian (Normal) Distribution

$$f_{xy}(x,y) = \frac{1}{2\pi \sigma_x \sigma_y \sqrt{1-\rho^2}} \exp\left[-\frac{1}{2}\frac{1}{1-\rho^2}\left(\hat{x}^2 - 2\rho\hat{x}\hat{y} + \hat{y}^2\right)\right]$$

$$\hat{x} = \frac{x-m_x}{\sigma_x} \qquad \hat{y} = \frac{y-m_y}{\sigma_y}$$

parameters: $m_x, m_y, \sigma_x, \sigma_y, \rho : |\rho| < 1$

See the accompanying plots of the 2D normal distribution.

B. Moments: Correlation and covariance

The expectation operator for 2D is now a double integral

$$E[x] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x \, f_{xy}(x,y)\, dx\, dy$$

$$E[y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y \, f_{xy}(x,y)\, dx\, dy$$

The 2nd moment is more interesting. Now, there are three of them:

$$\underbrace{E[x^2],\ E[y^2]}_{\text{2nd moment}},\ \underbrace{E[xy]}_{\text{correlation of } x \,\&\, y}$$

Centered moments are also useful:

$$\left.\begin{aligned} \mathrm{var}(x) &= E[(x-m_x)^2] \\ \mathrm{var}(y) &= E[(y-m_y)^2] \end{aligned}\right\} \text{variance}$$

$$\mathrm{cov}(x,y) = E[(x-m_x)(y-m_y)] \quad \text{covariance}$$

There is also a correlation coefficient:

$$\rho_{xy} = E\left[\left(\frac{x-m_x}{\sigma_x}\right)\left(\frac{y-m_y}{\sigma_y}\right)\right] = \frac{\mathrm{cov}(x,y)}{[\mathrm{var}(x)\,\mathrm{var}(y)]^{1/2}}$$

Uncorrelated, therefore, means $\rho_{xy} = 0$ or $\mathrm{cov}(x,y) = 0$

Aside: Mutually exclusive, independence, and correlation

Mutually exclusive implies $P(X \cap Y) = 0$. So, the joint cdf/pdf will give $P = 0$ for those values x and y are mutally exclusive.

$$P(X \text{ and } Y) = 0 \qquad P(X \text{ or } Y) = P(x) + P(y)$$

Independence implies $f_{xy}(x,y) = f_x(x)\, f_y(y)$. In this case the variables will always be uncorrelated. ← see below

$$P(X \text{ and } Y) = P(X)\, P(Y)$$

Proof:
$$\begin{aligned} \mathrm{cov}(x,y) &= E[(x - m_x)(y - m_y)] \\ &= E[x - m_x]\, E[y - m_y] \\ &= (E[x] - m_x)(E[y] - m_y) = 0 \end{aligned}$$

Uncorrelated does not imply independent. The covariance could be zero for other reasons.

C. Characteristic Function

The joint pdf has a characteristic function, just like the 1D version. Here, we need a bivariate Fourier transform.

$$\begin{aligned} \phi_{xy}(\omega, v) &= E\left[e^{i\omega x + ivy} \right] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{xy}(x,y)\, e^{i\omega x + ivy}\, dx\, dy \end{aligned}$$

D. Conditional Probability

The conditional probability density is defined as

$$f_{y|x}(y|x) = \frac{f_{xy}(x,y)}{f_x(x)}$$

← joint density

← marginal density

conditional density of y given x

$f_x(x)$ must also be non-zero

The definition above implies that

$$f_{xy}(x,y) = f_{y|x}(y|x)\, f_x(x)$$

can get joint density from conditional. Used something similar for Bayes theorem.

and a "law of total probability"

$$f_y(y) = \int_{-\infty}^{\infty} f(y|x)\, f_x(x)\, dx.$$

← can "integrate out" x

Independence means that

$$f_{y|x}(y|x) = f_y(y)$$

← f_y doesn't depend on x "drop the conditioning"

Using the above expression for the joint density gives

$$f_{xy}(x,y) = f_{y|x}(y|x)\, f_x(x) = f_y(y)\, f_x(x)$$

Mutually exclusive means that

$$f_{y|x}(y|x) = 0$$

(because $f_{xy} = 0$)

One can also define a conditional cdf

$$F_{y|x} = \int_{-\infty}^{y} f_{y|x}(t|x)\, dt$$

which is also equal to

$$F_{y|x} = \frac{\partial F_{xy}}{\partial y}.$$

← how does the joint distribution change with y at fixed x